

Listing of the Claims:

Claims 1-10 (Canceled)

Claim 11 (Previously amended): An automated method of assisting a user with the task of categorizing electronic documents into a collection, comprising the steps of:

classifying, with a classifier, a document to obtain a plurality of most likely categorical labels;

deriving a plurality of categorizations shortcuts from the plurality of most likely categorical labels;

displaying, to the user, a representation of the plurality of most likely categorical labels;

receiving , from the user, a selection of one or more of the most likely categorical labels representative of the document to be categorized within a collection;

labeling the document within the collection with one or more of the selected categorical labels; and

incrementally retraining the classifier to adapt to modifications of the collection, wherein the incremental retraining is performed using a lazy strategy for incrementally retraining the classifier.

Claim 12 (Canceled).

Claim 13 (Previously amended): The method of claim 11 wherein the classifying step comprises the step of classifying, upon receipt into data storage, the document to obtain the plurality of most likely categorical labels.

Claim 14 (Currently amended): The method of claim ~~12~~ 11 wherein the deriving step comprises the step of deriving, upon receipt of the document into data storage, categorization shortcuts from the plurality of most likely categorical labels.

Claim 15 (Currently amended): The method of claim ~~12~~ 11 wherein the deriving step comprises the step of labeling display buttons with the plurality of most likely categorical labels, and the displaying step comprises the step of displaying the labeled display buttons with the document.

Claim 16 (Currently amended): The method of claim ~~12~~ 11 wherein the deriving step comprises the step of creating an ordered set of the plurality of most likely categorical labels, and the displaying step comprises the step of displaying with the document the ordered set prepended to a standard ordered set of other categorical labels.

Claim 17 (Original): The method of claim 11 wherein the classifying step occurs substantially simultaneously with the displaying step.

Claim 18 (Original): The method of claim 11 wherein the classifying step comprises the step of classifying upon invocation by the user, the document to obtain the plurality of most likely categorical labels.

Claim 19 (Original): The method of claim 18 wherein the invocation comprises a selection by the user of a classify button.

Claim 20 (Original): The method of claim 11 wherein the labeling step comprises the step of storing the document in folders or locations of the collection corresponding to the one or more selected categorical labels.

Claim 21 (Original): The method of claim 11 further comprising the step of displaying a standard list of all categorical labels, wherein the receiving step comprises the step of receiving, from the user, data representative of one or more selected categorical labels from either the plurality of displayed categorization shortcuts or the standard list.

Claim 22 (Canceled).

Claim 23 (Previously amended): The method of claim 11 wherein the retraining step comprises the step of retraining the classifier in response to the labeling step.

Claim 24 (Previously amended): The method of claim 11 wherein the labeling step comprises the step of storing the document in folders or locations of the collection corresponding to the one or more selected categorical labels and the retraining step comprises the steps of:

receiving, from the user, addition data representative of an addition of a document into a tofolder; and

retraining the classifier in response to the addition data.

Claim 25 (Original): The method of claim 24 wherein the retraining step comprises the step of assigning, in the classifier, the added document to the tofolder.

Claim 26 (Original): The method of claim 25 further comprising the step of identifying excluded folders to be excluded from retraining and wherein the retraining step comprises the step of assigning, in the classifier, the added document when the tofolder is not one of the identified excluded folders.

Claim 27 (Previously amended): The method of claim 11 wherein the labeling step comprises the step of storing the document in folders or locations of the collection corresponding to the one or more selected categorical labels and the retraining step comprises the steps of:

receiving, from the user, deletion data representative of a removal of a document from a from folder; and

retraining the classifier in response to the deletion data.

Claim 28 (Original): The method of claim 27 wherein the retraining step comprises the step of unassigning, in the classifier, the removed document from the from folder in which it was categorized.

Claim 29 (Original): The method of claim 28 further comprising the step of identifying

excluded folders to be excluded from retraining and wherein the retraining step comprises the step of unassigning, in the classifier, the removed document when the from folder is not one of the identified excluded folders.

Claim 30 (Previously amended): The method of claim 11 wherein the labeling step comprises the step of storing the document in folders or locations of the collection corresponding to the one or more selected categorical labels and the retraining step comprises the steps of:

receiving, from the user, move data representative of a movement of a document from a source folder to a destination folder; and

retraining the classifier in response to the move data.

Claim 31 (Original): The method of claim 30 wherein the retraining step comprises the steps of:

unassigning, the classifier, the moved document from the source folder in which it was categorized; and

assigning, in the classifier, the moved document to the destination folder.

Claim 32 (Original): The method of claim 31 further comprising the step of identifying excluded folders to be excluded from retraining and wherein the retraining step comprises the steps of:

unassigning, in the classifier, the moved document when the source folder is not one of the identified excluded folders; and

assigning, in the classifier, the moved document when the destination folder is not one of the identified excluded folders.

Claim 33 (Previously amended): The method of claim 11 wherein the retraining step occurs immediately after a collection modification.

Claims 34-36 (Canceled).

Claim 37 (Original): The method of claim 20, wherein the classifying step comprises the steps of :

tokenizing the document into different tokens;
tallying a number of occurrences of each token in the document;
computing, for each folder, a token weight of each token;
comparing, for each token, the number of occurrences and the token weights;
creating a similarity score in response to the comparing step; and
identifying a subset of folders for which the similarity score is highest.

Claim 38 (Original): The method of claim 37 further comprising the step of removing, from the identified subset, all folders from which the similarity score is lower than a default or specified threshold.

Claim 39 (Original): The method of claim 37, wherein the computing step comprises the

step of computing the token counts of each token in each of the folders.

Claim 40 (Original): The method of claim 37 wherein the tokenizing step comprises the steps of:

separately tokenizing different portions of the document; and
labeling the tokens according to the different portions;

Claim 41 (Original): The method of claim 25 wherein the classifying step comprises the steps of:


tokenizing the document into different tokens;
tallying a number of occurrences of each token in the document;
retrieving, for each folder, a tokencount of each token;
computing, for each folder, a token weight of each token;
comparing, for each token, the number of occurrences and the token weights;
creating a similarity score in response to the comparing step; and
identifying a subset of folders for which the similarity score is highest, and

wherein the assigning step comprises the step of adding the number of occurrences of each token to the tokencount of the tofolder.

Claim 42 (Original): The method of claim 28 wherein the classifying step comprises the steps of :

tokenizing the document into different tokens;

tallying a number of occurrences of each token in the document;
retrieving, for each folder, a tokencount of each token;
computing, for each folder, a token weight of each token;
comparing, for each token, the number of occurrences and the token weights;
creating a similarity score in response to the comparing step; and
identifying a subset of folders for which the similarity score is highest, and
wherein the unassigning step comprises the step of subtracting the number of occurrences of each token to the tokencount of the tofolder.




Claim 43 (Original): The method of claim 31 wherein the classifying steps comprises the steps of:

tokenizing the document into different tokens;
tallying a number of occurrences of each token in the document;
retrieving, for each folder, a tokencount of each token;
computing, for each folder, a token weight of each token;
comparing, for each token, the number of occurrences and the token weights;
creating a similarity score in response to the comparing step; and
identifying a subset of folders for which the similarity score is highest, and
wherein the unassigning step comprises the step of subtracting the number of occurrences of each token from the tokencount of the source folder, and the assigning step comprises the step of adding the number of occurrences of each token to the tokencount of the destination folder.

Claim 44 (Original): The method of claim 1, further comprising the step of training the classifier from scratch with a pre-existing collection of categorized documents.

Claim 45 (Original): The method of claim 44 wherein the labeling step comprises the step of storing the document in folders or locations of the collection corresponding to the one or more selected categorical labels and the training step comprises the step of assigning, in the classifier, each of the pre-existing documents to a folder in which it is categorized.

 Claim 46 (Original): The method of claim 45 wherein the classifying step comprises the steps of:

tokenizing the document into different tokens;

tallying a number of occurrences of each token in the document;

retrieving, for each folder, a tokencount of each token;

computing, for each folder, a token weight of each token;

comparing, for each token, the number of occurrences and the token weight;

creating a similarity score in response to the comparing step; and

identifying a subset of folders for which the similarity score is highest, and

wherein the assigning step comprises the step of adding the number of occurrences of each token to the tokencount of the tofolder.

Claim 47 (Original): The method of claim 45 further comprising the step of identifying excluded folders to be excluded from training and wherein the training step comprises the step of

assigning, in the classifier, each of the pre-existing documents, except those in the identified excluded folders.

Claim 48 (Original): The method of claim 11 wherein the labeling step comprises the step of storing the document in folders or locations of the collection corresponding to the one or more selected categorical labels and the re-training step comprises the steps of:

determining a time of a last step of retraining; and

retraining the classifier on each folder which was modified after the determined time.

Claim 49 (Previously amended): The method of claim 11 wherein labeling step comprises the step of storing the document in folders or locations of the collection corresponding to the one or more selected categorical labels, the method further comprising the step of training the classifier from scratch with a pre-existing collection of categorized documents, wherein the retraining step comprises the steps of:

determining a time of the step of training or a last step of re-training; and

retraining the classifier on each folder which was modified after the determined time.

Claim 50 (Original): The method of claim 11 wherein the classifying step uses the TF-IDF principle.

Claim 51 (Original): The method of claim 11 wherein the electronic document is an e-mail message.

Claim 52 (Original): The method of claim 11 wherein the electronic document is web page and the collection is a collection of bookmarks.

Claim 53 (Original): The method of claim 41 wherein the electronic document is web page and the collection is a collection of bookmarks, the method further comprising the step of storing, for each web page, a pagetokencount matching the tallied number of occurrences of each token.

Claim 54 (Original): The method of claim 42 wherein the electronic document is a web page and the collection is a collection of bookmarks, the method further comprising the step of storing, for each web page, a pagetokencount matching the tallied number of occurrences of each token, wherein the unassigning step comprises the step of subtracting the pagetokencount from the tokencount of the fromfolder.

Claim 55 (Original): The method of claim 43 wherein the electronic document is a web page and the collection is a collection of bookmarks, the method further comprising the step of storing, for each web page, a pagetokencount matching the tallied number of occurrences of each token, wherein the unassigning step comprises the step of subtracting the pagetokencount from the tokencount of the fromfolder.

Claim 56 (Original): The method of claim 11 wherein the electronic document is a multimedia document.

Claim 57 (Original): The method of claim 56 wherein the multimedia document is an image file, a video file or an audio file.

Claim 58 (Original): The method of claim 56 wherein the multimedia document combines any combination of text, an image file, a video file and an audio file.

Claim 59 (Original): The method of claim 57 wherein the multimedia document further includes text.

Claim 60 (Original): The method of claim 11 wherein the electronic document comprises data sets that are not viewable in their entirety, but can be categorized in response to some presentation to the user.

61. (Previously amended): A program storage device, readable by a machine, tangibly embodying a program of instructions executable by the machine to perform method steps for automatically assisting a user with the task of categorizing an electronic document into a collection, the method comprising the steps of:

classifying, with a classifier, a document to obtain a plurality of most likely categorical

labels;

deriving a plurality of categorizations shortcuts from the plurality of most likely categorical labels;

displaying, to the user, a representation of the plurality of most likely categorical labels;

receiving , from the user, a selection of one or more of the most likely categorical labels representative of the document to be categorized within a collection;

labeling the document within the collection with one or more of the selected categorical labels; and

incrementally retraining the classifier to adapt to modifications of the collection, wherein the incremental retraining is performed using a lazy strategy for incrementally retraining the classifier.

Claim 62 (Currently amended): The method of claim 11 wherein:

the step of displaying most likely categorical labels to the user further comprises displaying through a graphical user interface a plurality of virtual category buttons, each labeled with one of the categorical labels; ~~and~~

the step of receiving representative data from the user comprises receiving information that one of the category buttons has been selected by the user by clicking on the category button; and

the step of labeling the document comprises virtually moving the document to a virtual file folder unique to the categorical label selected, and occurs upon the clicking of the selected category button without need for any other activity by the user.

Claim 63 (Currently amended): The program storage device of claim 61 wherein:
the step of displaying most likely categorical labels to the user further comprises
displaying through a graphical user interface a plurality of virtual category buttons, each labeled
with one of the categorical labels; ~~and~~

the step of receiving representative data from the user comprises receiving information
that one of the category buttons has been selected by the user by clicking on the category button;
and

the step of labeling the document comprises virtually moving the document to virtual file
folder unique to the categorical label selected, and occurs upon the clicking of the selected
category button without need for any other activity by the user.

Claim 64 (Previously added): The automated method of claim 11, wherein the
incrementally retraining is performed using a lazy strategy for incrementally retraining the
classifier comprises:

deferring retraining of the classifier;

performing bookkeeping operations whenever messages are added to folders, removed
from folders, or moved from one folder to another; and

automatically triggering retraining of the classifier by a predetermined criteria, wherein
the retraining the classifier by processing any updates that have been deferred.

Claim 65 (Previously added): The automatic method of claim 64, wherein the

predetermined criteria for automatically performing the retraining step is a fixed amount of time that has elapsed since a last retraining step has been performed.

Claim 66 (Previously added): The automatic method of claim 64, wherein the predetermined criteria for automatically performing the retraining step is when a threshold number of documents have been added, deleted, or moved in the collection and any combination thereof.

Claim 67 (Currently amended): The automatic method of claim 64, wherein the predetermined criteria for automatically performing the retraining step is when the system has reached an idle state or is either updating the classifier or performing bookkeeping operations whenever messages are added to folders, removed from folders, or moved from one folder to another.

Claim 68 (Previously added): The automatic method of claim 11, wherein incremental retraining the classifier includes an instant strategy for incrementally retraining the classifier.
